Universality of the Distribution Functions of Random Matrix Theory

Craig A. Tracy
Department of Mathematics
Institute of Theoretical Dynamics
University of California
Davis, CA 95616, USA

Harold Widom
Department of Mathematics
University of California
Santa Cruz, CA 95064, USA

August 31, 1999

1 Random Matrix Models

In probability theory and statistics a common first approximation to many random processes is a sequence $X_1, X_2, X_3, \ldots$ of independent and identically distributed (iid) random variables. Let F denote their common distribution. To motivate the material below, we take these random variables and construct a particularly simple $N \times N$ random matrix,

$$\mathrm{diag}\left(X_1(\omega), X_2(\omega), \cdots, X_N(\omega)\right).$$

The order statistics are the eigenvalues ordered

$$\lambda_1 \le \lambda_2 \le \cdots \le \lambda_N,$$

and the distribution of the largest eigenvalue, $\lambda_{\max}(N) = \lambda_N$, is

$$\begin{aligned} \mathrm{Prob}\left(\lambda_{\max}(N) \le x\right) &= \mathrm{Prob}\left(X_1 \le x, \ldots, X_N \le x\right) \\ &= F(x)^N. \end{aligned}$$

Since the distribution F is arbitrary, we observe that so too is the distribution of the largest eigenvalue of a $N \times N$ random matrix. However, one is really interested in limiting laws as $N \to \infty$. That is, we ask if there exist constants a_N and b_N such that

$$\frac{\lambda_{\max}(N) - a_N}{b_N} \tag{1.1}$$

converges in distribution to a nontrivial limiting distribution function G. In the present situation a complete answer is provided by

Theorem: If (1.1) converges in distribution to some nontrivial distribution function G, then G belongs to one of the following forms:

1. $e^{-e^{-x}}$ with support $\mathbf{R}$.
2. e^{-1/x^α} with support $[0,\infty)$ and $\alpha > 0$.
3. $e^{-(-x)^\alpha}$ with support $(-\infty,0]$ and $\alpha > 0$.

This theorem is a model for the type of results we want for nondiagonal random matrices.

A random matrix model is a probability space $(\Omega, \mathcal{P}, \mathcal{F})$ where Ω is a set of matrices. Here are some examples

- Circular Unitary Ensemble (CUE, $\beta = 2$)
  - $\Omega = \mathcal{U}(N) = N \times N$ unitary matrices.
  - $\mathcal{P}$= Haar measure.
- Gaussian Orthogonal Ensemble (GOE, $\beta = 1$)
  - $\Omega = N \times N$ real symmetric matrices.
  - $\mathcal{P}$ = unique[1] measure that is invariant under orthogonal transformations and the matrix elements (say on and above the diagonal) are iid random variables.
- Gaussian Unitary Ensemble (GUE, $\beta = 2$)
  - $\Omega = N \times N$ hermitian matrices.
  - $\mathcal{P}$= unique measure that is invariant under unitary transformations and the real and imaginary matrix elements (say on and above the diagonal) are iid random variables.
- Gaussian Symplectic Ensemble (GSE, $\beta = 4$)
  - $\Omega = 2N \times 2N$ Hermitian self-dual matrices.[2]
  - $\mathcal{P}$= unique measure that is invariant under symplectic transformations and the real and imaginary matrix elements (say on and above the diagonal) are iid random variables.

[1]Uniqueness is up to centering and a normalization of the variance.

[2]Identify the $2N \times 2N$ matrix with the $N \times N$ matrix whose entries are quaternions. If the quaternion matrix elements satisfy $\overline{M}_{ji} = M_{ij}$ where the bar is quaternion conjugation, then the $2N \times 2N$ matrix is called Hermitian self-dual. Each eigenvalue of a Hermitian self-dual matrix has multiplicity two.

Expected values of random variables $f : \Omega \to \mathbf{C}$ are computed from the usual formula

$$E_\Omega(f) = \int_\Omega f(M)\, d\mathcal{P}(M).$$

If $f(M)$ depends only on the eigenvalues of $M \in \Omega$, then one can be more explicit:

- CUE (Weyl's Formula)

$$E_{\mathcal{U}(N)}(f) = \frac{1}{N!(2\pi)^N} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} f(\theta_1, \ldots, \theta_N) \prod_{1 \le \mu < \nu \le N} \left| \Delta(e^{i\theta_1}, \ldots, e^{i\theta_N}) \right|^2 d\theta_1 \cdots d\theta_N,$$

- Gaussian Ensembles ($\beta = 1, 2, 4$):

$$E_{N\beta}(f) = c_{N\beta} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_N)\, |\Delta(x_1, \ldots, x_N)|^\beta\, e^{-\frac{\beta}{2}\sum x_j^2}\, dx_1 \cdots dx_N,$$

where $c_{N\beta}$ is chosen so that $E_{N\beta}(1) = 1$ and $\Delta(x_1, \ldots, x_N) = \prod_{1\le i<j\le N}(x_i - x_j)$. The factor $e^{-\frac{\beta}{2}\sum x_j^2}$ explains the choice of the word "gaussian" in the names of these ensembles. A commonly studied generalization of these gaussian measures is to replace the sum of quadratic terms appearing in the exponential with $\sum V(x_i)$ where V is, say, a polynomial (with the obvious restrictions to make the measure well-defined).

Choosing $f = \prod_i \left(1 - \chi_J(x_i)\right)$, χ_J the characteristic function of a set $J \subset \mathbf{R}$, we get the important quantity[3]

$$E_{N\beta}(f) = E_{N\beta}(0; J) := \text{probability no eigenvalues lie in } J,$$

and in the particular case $J = (t, \infty)$

$$F_{N\beta}(t) := \text{Prob}\,(\lambda_{\max} \le t) = E_{N\beta}(0, J).$$

The level spacing distribution[4] is expressible in terms of the mixed second partial derivative of $E_{N\beta}(0; (a, b))$ with respect to the endpoints a and b.

2 Fredholm Determinant Representations

Though $E_{N\beta}(0; J)$ are explicit N-dimensional integrals, these expressions are not so useful in establishing limiting laws as $N \to \infty$. What turned out to be very useful are Fredholm

[3]This quantity has an obvious extension to other random matrix models.

[4]Let the eigenvalues be ordered. The conditional probability that given an eigenvalue at a, the next one lies between s and $s + ds$ is called the level spacing density.

determinant representations for $E_{N\beta}(0;J)$. In 1961 M. Gaudin proved for $\beta = 2$ (using the newly developed orthogonal polynomial method of M. L. Mehta) that $E_{N2}(0;J) = \det(I - K_{N2})$ where K_{N2} is an integral operator acting on J whose kernel is of the form

$$\frac{\varphi(x)\psi(y) - \psi(x)\varphi(y)}{x-y}, \tag{2.1}$$

with $\varphi(x) = c_N e^{-x^2/2} H_N(x)$, $\psi(x) = c_N e^{-x^2/2} H_{N-1}(x)$, and $H_j(x)$ are the Hermite polynomials.[5] For $\beta = 1$ or 4, generalizing F. J. Dyson's 1970 analysis of the n-point correlations for the circular ensembles, it follows from work by Mehta the following year that the square of $E_{N\beta}(0;J)$ again equals a Fredholm determinant, $\det(I - K_{N\beta})$, but now the kernel of $K_{N\beta}$ is a 2×2 matrix.[6]

3 Scaling Limits (Limiting Laws)

3.1 Bulk Scaling Limit

Let $\rho_N(x)$ denote the density of eigenvalues at x and pick a point x_0, independent of N with $\rho_N(x_0) > 0$. We scale distances so that resulting density is one at x_0, $\xi := \rho_N(x_0)\,(x - x_0)$, and we call the limit

$$N \to \infty, x \to x_0, \text{ such that } \xi \text{ is fixed},$$

the bulk scaling limit. For $\beta = 2$,

$$E_{N2}(0;J) \to E_2(0;J) = \det\left(I - K_2\right)$$

where the integral operator K_2 (acting on $L^2(J)$) has as its kernel (the sine kernel)

$$\frac{1}{\pi}\frac{\sin \pi(\xi - \xi')}{\xi - \xi'}.$$

(We use the same symbol J to denote the scaled set J.) Furthermore,

$$p_2(s) = -\frac{d^2}{ds^2} E_2\left(0;(0,s)\right)$$

is the (limiting) level-spacing density for GUE; known as the Gaudin distribution.[7] We observe that the limiting kernel is translationally invariant and independent of x_0.

[5] For the random matrix models corresponding to general potential V, $\varphi(x) = c_N e^{-V(x)/2} p_N(x)$ and $\psi(x) = c_N e^{-V(x)/2} p_{N-1}(x)$ where $p_j(x)$ are the orthogonal polynomials associated with weight function $w(x) = e^{-V(x)}$. It is in this generalization that we see the close relation between the general theory of orthogonal polynomials and random matrix theory.

[6] See [37] for elementary proofs of these facts.

[7] For the analogous $\beta = 1, 4$ results, see, e.g., [27] or [37].

3.2 Edge Scaling Limit

In the gaussian ensembles, the density decays exponentially fast around $2\sigma\sqrt{N}$; perhaps surprisingly, it is also the case that

$$\lim_{N\to\infty} \frac{\lambda_{\max}(N)}{\sqrt{N}} = 2\sigma,\ a.s. \tag{3.1}$$

where σ is the standard deviation of the off-diagonal matrix elements. (In the normalization we've adopted, $\sigma = 1/\sqrt{2}$.) If we introduce the scaled random variable $\hat{\lambda}$ through

$$\lambda_{\max} = 2\sigma\sqrt{N} + \frac{\sigma\hat{\lambda}}{N^{1/6}},$$

then

$$\mathrm{Prob}\,(\lambda_{\max} \le t) = \mathrm{Prob}\left(\hat{\lambda} \le s\right) \to F_\beta(s) \ \text{ as } N \to \infty,$$

where $t = 2\sigma\sqrt{N} + \sigma s/N^{1/6}$. For $\beta = 2$,

$$F_2(s) = \det(I - K_{\mathrm{Airy}}),$$

where K_{Airy} has kernel of the form (2.1) with $\varphi(x) = \mathrm{Ai}(x)$, $\psi(x) = \mathrm{Ai}'(x)$ and $J = (s, \infty)$. (See, e.g., [37] for the $\beta = 1, 4$ results.)

4 Connections with Integrable Systems

4.1 Bulk Scaling Limit

In 1980 M. Jimbo, T. Miwa, Y. Môri, and M. Sato [21] expressed the Fredholm determinant of the sine kernel in terms of a solution to a certain system of integrable differential equations.[8] In the simplest case of a single interval, $J = (0, s)$, the differential equation is a particular case of Painlevé V (P_V)[9] and the Fredholm determinant is given by

$$\det(I - \lambda K_2) = \exp\left(\int_0^{\pi s} \frac{\sigma(x;\lambda)}{x}\,dx\right),$$

$$(x\sigma'')^2 + 4\,(x\sigma' - \sigma)\left(x\sigma' - \sigma + (\sigma')^2\right) = 0,$$

$$\sigma(x;\lambda) \sim -\frac{\lambda}{\pi}\,x, \ \text{ as } x \to 0.$$

For $\beta = 1, 4$ and $J = (0, s)$, $E_\beta(0; (0, s))$ can also be expressed in terms of the same function $\sigma(x; 1)$. A down-to-earth application of these Painlevé representations (and using the known asymtptotics of $\sigma(x; 1)$) is that one can easily produce graphs of the level spacing densities $p_\beta(s)$.[10]

[8] A simplified proof of their results can be found in [38].

[9] The differential equation below is the sigma representation of P_V.

[10] Without the Painlevé representations, the numerical evaluation of the Fredholm determinants is quite involved.

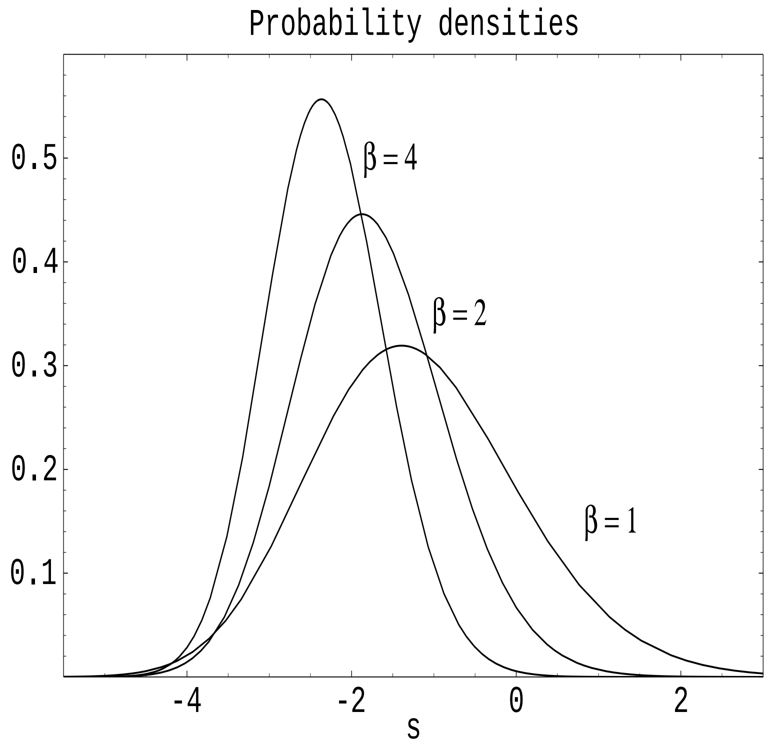

Figure 1: Densities for the scaled largest eigenvalues, $f_\beta(s)$.

4.2 Edge Scaling Limit

The limiting distributions (edge scaling) of the largest eigenvalue, $F_\beta(s)$, can also be expressed in terms of Painlevé functions—this time P_{II} [39, 40]:

$$F_1(s)^2 = \exp\left(-\int_s^\infty q(x)\,dx\right) F_2(s), \tag{4.1}$$

$$F_2(s) = \exp\left(-\int_s^\infty (x-s)q(x)^2\,dx\right), \tag{4.2}$$

$$F_4(s/\sqrt{2})^2 = \cosh^2\left(\frac{1}{2}\int_s^\infty q(x)\,dx\right) F_2(s), \tag{4.3}$$

where q satisfies the Painlevé II equation

$$q'' = xq + 2q^3$$

with boundary condition $q(x) \sim \mathrm{Ai}(x)$ as $x \to \infty$.[11] The graphs of the densities $f_\beta(s) = dF_\beta(s)/ds$ are in Figure 1.

[11] That a solution q exists and is unique follows from the representation of the Fredholm determinant in terms of it. Independent proofs of this, as well as the asymptotics as $x \to -\infty$, were given by S. Hastings and J. McLeod, P. Clarkson and McLeod and by P. Deift and X. Zhou.

4.3 Generalizations

Both the sine kernel and the Airy kernel are of the form (2.1). Kernels of this form arise in many problems in integrable systems; indeed, so much so that A. Its, A. Izergin, V. Korepin and V. Slavnov [20] in 1990 initiated a general analysis of these kernels. The following theorem [41], which applies to a wide class of $\beta = 2$ random matrix ensembles, gives the general situation:

Theorem: Let $J = \bigcup_{j=1}^{m}(a_{2j-1} - a_{2j})$ be a union of open intervals. Define $\tau(a) = \det(I - K)$ where K is an integral operator acting on $L^2(J)$ whose kernel is of the form (2.1) where φ and ψ are assumed to satisfy

$$\frac{d}{dx}\begin{pmatrix} \varphi \\ \psi \end{pmatrix} = \Omega(x)\begin{pmatrix} \varphi \\ \psi \end{pmatrix}$$

with $\Omega(x)$ a 2×2 matrix with zero trace and rational entries in x, then $\frac{\partial}{\partial a_j} \log \det(I - K)$ are expressible polynomially in terms of solutions to a total system of partial differential equations (a_j are the independent variables). The differential equations are given explicitly in terms of the coefficients of the rational functions appearing in $\Omega(x)$.

4.4 Historical Comments

The first connection between Toeplitz/Fredholm determinants and Painlevé functions was established in 1973–77 in work of T. T. Wu, B. M. McCoy, E. Barouch and the first author concerning the scaling limit of the 2-point functions of the 2D Ising model of statistical mechanics. The Painlevé function that arose was P_{III}. This work was subsequently generalized by Sato, Miwa and Jimbo to n-point functions and, more generally, holonomic quantum fields. The Kyoto School then took up the problem of the density matrix of the impenetrable Bose gas and it was in this context that they discovered that the Fredholm determinant of the sine kernel is related to P_V.

A crucial simplification of the Kyoto School work, as it applies to random matrix theory, was made by Mehta in 1992 [28]. This last work inspired the commutator methods introduced by the present authors in the period 1993–96. Since then both Riemann-Hilbert methods of Deift, Its, Zhou and others (see, e.g. [15]); and Virasoro methods of M. Adler, T. Shiota, P. van Moerbeke, and others (see, e.g. [1]), have played an increasingly important role in the development of random matrix theory. The connection of these methods with the isomonodromy method has been clarified by J. Palmer [32] and J. Harnad [19].

Space does not permit us to discuss the interesting connections between random matrices and Szegö type limit theorems. See E. Basor [7] for connections with linear statistics and the review papers [8, 42] for some related historical comments.

5 Universality

5.1 Universality of Gaussian Ensembles in Random Matrix Models

5.1.1 Invariant Measures, $\beta = 2$

As briefly mentioned above, a widely studied class of random matrix models is defined by the replacement of the gaussian potential, x^2, by the general potential $V(x)$. For the weight functions most studied, the parameter N is put into the exponent so that the weight function becomes $e^{-NV(x)}$. For different V's, the limiting density $\rho_V(x)$ can be quite different. It may be supported on many distinct intervals, and it may vanish at interior points of its support. In the gaussian case, the limiting density is the Wigner semicircle law: $\rho_W(x) = \frac{2}{\pi}\sqrt{1-x^2}$. Heuristic arguments suggest that the behavior exhibited by the Wigner law—that ρ is positive on the interior of its support and vanishes like a square root at endpoints—is the typical behavior for ρ_V. The bulk scaling limit and edge scaling limit are defined in analogous ways to the gaussian cases. To establish universality of these scaling limits, one must show (for $\beta = 2$ ensembles) that the scaled kernels approach the sine kernel and the Airy kernel, respectively. The potential $V(x) = \frac{t}{2}x^2 + \frac{g}{4}x^4$ ($g > 0$, $t < 0$) is an example of a "two interval" potential. Indeed, for this important example P. Bleher and A. Its [9] proved precisely this statement of universality. (See their paper for related work in the orthogonal polynomial literature as well as the physics literature.) Recently, building on work of [16], A. Kuijlaars and K. McLaughlin [26] have shown this behavior is generic for real analytic V satisfying $\lim_{x\to\infty} V(x)/\log|x| = +\infty$.

In the physics literature, M. Bowick, E. Brézin [12] and others have argued (for $\beta = 2$ ensembles) that if ρ_V vanishes faster than a square root, then the corresponding edge scaling limit will result in nonAiry universality classes. The resulting new kernels will have form (2.1) and the theory developed in [41] will apply, but there remains much to be understood in these cases.

For $\beta = 1, 4$, the situation is more complicated due to the structure of K_V [37, 45], and the "universality" theorems are not so general.

5.2 Noninvariant Measures: Wigner Ensemble

The Wigner ensembles are defined by requiring that the matrix elements on and above the diagonal in either the real symmetric case or the complex hermitian case are independent and identically distributed random variables. It is only in the case when the distribution is gaussian is the measure invariant. One usually assumes, as we do here, that all moments of the common distribution function exist. It was Wigner himself who showed that the limiting density of states is the Wigner semicircle. Subsequently several authors—culminating in a theorem by Z. Bai and Y. Lin clarifying which moments need exist—showed (3.1) continues to hold for the Wigner ensembles.

It should be noted that because the measure is noninvariant, the nongaussian Wigner ensembles do not, as far as we understand, have Fredholm determinant representation for their distribu-

tion functions. This means, for one, that the methods of integrable systems are not directly applicable to Wigner ensembles. It is therefore particularly important, as A. Soshnikov [36] recently proved, that in the edge scaling limit the Wigner ensembles are in the same universality class as the gaussian models. In particular, the limiting distribution of the scaled largest eigenvalue is given by $F_1(s)$ for real symmetric Wigner matrices and by $F_2(s)$ for complex hermitian Wigner matrices.

5.3 Examples from Physics

A second type of universality, and the one first envisioned by Wigner in the context of nuclear physics, asserts in Wigner's words [46]

> Let me say only one more word. It is very likely that the curve in Figure I [an approximate graph of $p_1(s)$] is a universal function. In other words, it doesn't depend on the details of the model with which you are working.

The modern version of this asserts that for a classical, "fully" chaotic Hamiltonian the corresponding quantum system has a level spacing distribution equal to $p_\beta(s)$ in the bulk. (The symmetry class determines which ensemble.) This quantum chaos conjecture, due to O. Bohigas, M. Giannoni and C. Schmit [10], has been a guiding principle for much subsequent work, though it is the authors' understanding that it remains a conjecture. A particularly nice numerical example supporting this conjecture is M. Robnik's work [35] on chaotic billards. The reader is referred to the recent review article [18] for further numerical examples that support this conjecture. It should be noted that there are examples from number theory where the conjecture fails. Thus, as it has been said, the conjecture is undoubtedly true except where it is demonstratively false.

5.3.1 Aperiodic Tiling Adjacency Matrix

The discovery of quasicrystals has made the study of statistical mechanical models whose underlying lattice is quasiperiodic of considerable interest to physicists. In particular, in order to understand transport properties, tight binding models have been defined on various quasiperiodic lattices. One such study by Zhong *et al.* [47] defined a simplified tight binding model for the octagonal tiling of Ammann and Beenker. This quasiperiodic tiling consists of squares and rhombi with all edges of equal lengths (see Figure 2) and has a D_8 symmetry around the central vertex. On this tiling the authors take as their Hamiltonian the adjacency matrix for the graph with free boundary conditions. The largest lattice they consider has 157,369 vertices. The matrix splits into ten blocks according to the irreducible representations of the dihedral group D_8. For each of these ten independent subspectra, they compare the empirical distribution of the normalized spacings of the consecutive eigenvalues with the GOE level spacing density $p_1(s)$. In Figure 2 we have reproduced a portion of their data for one such subspectrum together with p_1.

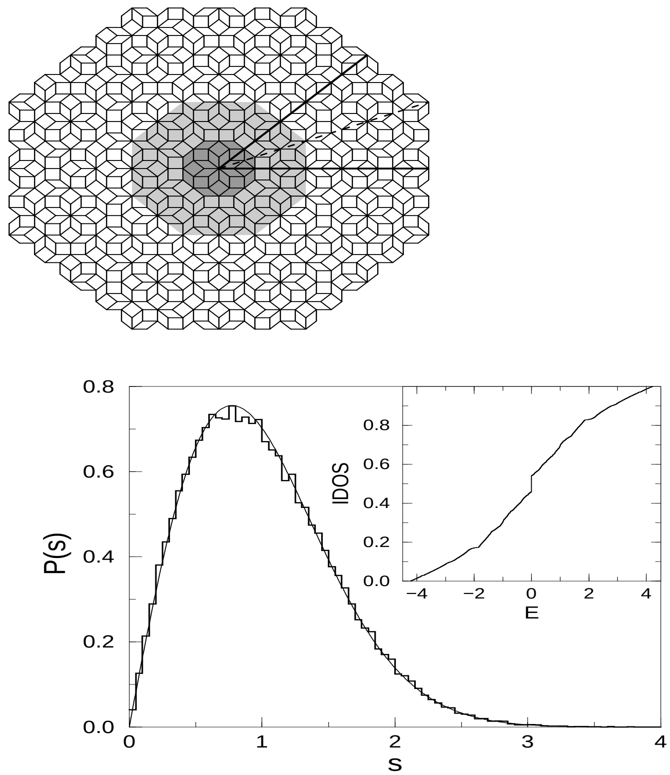

Figure 2: Data for nearest neighbor normalized spacings of eigenvalues of the adjacency matrix for a quasiperiodic octagonal tiling are plotted together with the GOE level spacing density $p_1(s)$. Data are from one independent subspectrum of a D_8-symmetric octagonal patch of a tiling with 157,369 vertices. Courtesy of Zhong *et al.* [47].

5.4 Spacings of the Consecutive Zeros of Zeta Functions

Perhaps the most surprising appearance of the distributions of random matrix theory is in number theory. Analytical work by H. Montgomery and extensive numerical calculations by A. Odlyzko on the zeros of the Riemann zeta function have given convincing evidence that the normalized consecutive spacings follow the Gaudin distribution, see Figure 3. Recent results of Z. Rudnick and P. Sarnak are also compatible with the belief that the distribution of the spacings between zeros, not only of the Riemann zeta function, but also of quite general automorphic L-functions over $\mathbf{Q}$, are all given by this Montgomery-Odlyzko Law. In their landmark book [25], N. Katz and P. Sarnak establish the Montgomery-Odlyzko Law for wide classes of zeta and L-functions over finite fields.

5.5 Random Matrix Theory and Combinatorics

The last decade has seen a flurry of activity centering around connections between combinatorial probability of the Robinson-Schensted-Knuth (RSK) type on the one hand and random matrices and integrable systems on the other. From the point of view of probability theory, the quite surprising feature of these developments is that the methods came from Toeplitz determinants, integrable differential equations of the Painlevé type and the closely related Riemann-Hilbert

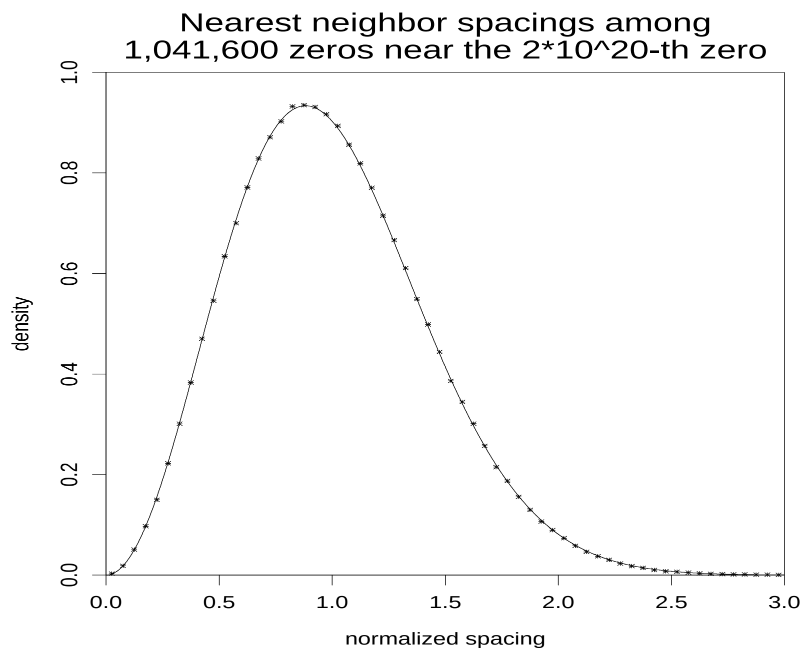

Figure 3: Data for nearest neighbor spacings among 1,041,600 zeros of the Riemann zeta function near the 2×10^{20}-th zero are plotted together with the GUE spacing density. Courtesy of Andrew Odlyzko [29].

techniques as they were applied and refined in random matrix theory. Using these techniques new, and apparently quite universal, limiting laws have been discovered. The earliest signs of these connections can be found in the work of A. Regev [34] and I. Gessel [17]. Here, however, we introduce this subject by examining a certain card game of D. Aldous and P. Diaconis [3], called patience sorting.

5.5.1 Patience Sorting and Random Permutations

Our deck of cards is labeled $\{1, 2, \dots, N\}$ and we order the cards with their natural ordering. Shuffle the deck of cards and

- Turn over the first card.
- Turn over the second card. If it is of higher rank, start a new pile to the right of the first card. Otherwise place the second card on top of the first card.
- Turn over the third card. If it is of higher rank than either the first or the second card, start a new pile to the right of the second card. Otherwise place the third card on top of the card of higher rank. If both first and second are of higher rank, place the third card on the smaller ranked card. (That is, play cards as far as possible to the left.)

- Continue playing the game, playing cards as far left as possible, until all the cards are turned over.

The object of the game is to end with a small number of piles. Let $\ell_N(\sigma)$ equal the number of piles at the end of the game where we started with deck $\sigma = \{i_1, i_2, \ldots, i_N\}$. Clearly, $1 \leq \ell_N(\sigma) \leq N$, but what are some typical values for a shuffled deck? Starting each time with a newly shuffled deck of $N = 52$ cards, the computer played patience sorting 100,000 times. Here are the statistics for ℓ_{52}:

- Mean=11.56 (11.00).
- Standard Deviation=1.37 (1.74)
- Skewness=0.33 (0.22)
- Kurtosis Excess =0.16 (0.09)
- Sample Range = 7 to 19 (Probability 0.993)

where the numbers in parentheses are the asymptotic predictions (as the number of cards tends to infinity) of the theory of J. Baik, P. Deift and K. Johansson [4] to be described below.

A shuffled deck of cards, $\sigma = \{i_1, i_2, \ldots, i_N\}$, is a permutation of $\{1, 2, \ldots, n\}$, and so we think of the shuffled deck as a random permutation. A moment's reflection will convince the reader that $\ell_N(\sigma)$ is equal to the length of the longest increasing subsequence in the permutation σ. As a problem in random permutations, determining the asymptotics of $E(\ell_N)$ as $N \to \infty$ is called Ulam's Problem. In the 1970's A. Vershik and S. Kerov and independently B. Logan and L. Shepp showed $E(\ell_N) \sim 2\sqrt{N}$ with important earlier work by J. Hammersley. Hammersley's analysis introduced a certain interacting particle system interpretation. This was developed by Aldous and Diaconis [2] who in 1995 gave a "soft" proof of this result using hydrodynamic scaling arguments from interacting particle theory.

Introducing the exponential generating function

$$\sum_{N \geq 0} \mathrm{Prob}(\ell_N \leq n) \, \frac{t^N}{N!} \,,$$

Gessel showed that it is equal to $D_n(t)$, the determinant of the $n \times n$ Toeplitz determinant with symbol $e^{\sqrt{t}(z+z^{-1})}$. (Recall that the i, j entry of a Toeplitz matrix equals the $i-j$ Fourier coefficient of its symbol.) It is in this work of Gessel and subsequent work of Odlyzko *et al.* [30] and E. Rains [33], that the methods of random matrix theory first appear in RSK type problems.[12]

[12]Gessel [17] does not mention random matrices, but in light of well-known formulas in random matrix theory relating Toeplitz determinants to expectations over the unitary group, we believe it is fair to say that the connection with random matrix theory begins with this discovery. See, however, Regev [34].

Starting with this Toeplitz determinant representation, Baik, Deift and Johansson [4], using the steepest descent method for Riemann-Hilbert problems [14], derived a delicate asymptotic formula for $D_n(t)$ which we now describe. Introduce another parameter s and suppose that n and t are related by $n = [2t^{1/2} + st^{1/6}]$. Then as $t \to \infty$ with s fixed one has

$$\lim_{t\to\infty} e^{-t} D_n(t) = F_2(s)$$

where $F_2(s)$ is the distribution function (4.2). Using a dePoissonization lemma due to Johansson [22], these asymptotics led Baik, Deift and Johansson to the limiting law

$$\lim_{N\to\infty} \text{Prob}\left(\frac{\ell_N - 2\sqrt{N}}{N^{1/6}} < s\right) = F_2(s).$$

Since the work of Baik, Deift and Johansson, several groups have extended this connection between RSK type combinatorics and the distribution functions of random matrix theory. The aforementioned result is equivalent to the determination of the limiting distribution of the number of boxes in the first row in the RSK correspondence $\sigma \leftrightarrow (P, Q)$. In [5] the same authors show that the limiting distribution of the number of boxes in the *second* row is (when centered and normalized) distributed as the *second* largest scaled eigenvalue in GUE [37]. They then conjectured that this correspondence extends to all rows. This conjecture was recently proved by A. Okounkov [31] using topological methods and by A. Borodin, A. Okounkov and G. Olshanski [13] and Johansson [24] using analytical methods.

Placing restrictions on the permutations σ (that they be fixed point free and involutions), Baik and Rains [6] have shown that the limiting laws for the length of the longest increasing/decreasing subsequence are now the limiting distributions F_1 and F_4 for the scaled largest eigenvalue in GOE and GSE, see (4.1) and (4.3). Generalizing to signed permutations and colored permutations the present authors and Borodin [43, 11] showed that the distribution functions of the length of the longest increasing subsequence involve the same F_2.

Johansson [23] showed that the shape fluctuations of a certain random growth model, again appropriately scaled, converges in distribution to F_2. (This random growth model is intimately related to certain randomly growing Young diagrams.) In subsequent work, Johansson [24] showed that the fluctuations in certain random tiling problems (related to the Artic Circle Theorem) are again described by F_2. Finally, Johansson [24] and the present authors [44] have considered analogous problems for random *words* and have discovered various random matrix theory connections.

Acknowledgments The authors have benefited from conversations with A. Its and it is a pleasure to acknowledge this. The first author thanks J. Harnad and P. Winternitz for their invitation to speak at the workshop *Integrable Systems: From Classical to Quantum*. This work was supported, in part, by the National Science Foundation through grants DMS–9802122 and DMS–9732687.

References

[1] M. Adler, T. Shiota, and P. van Moerbeke, Random matrices, Virasoro algebras, and noncommutative KP, *Duke Math. J.* **94** (1998) 379–431.

[2] D. Aldous and P. Diaconis, Hammersley's interacting particle process and longest increasing subsequences, *Probab. Theory Related Fields* **103** (1995) 199–213.

[3] D. Aldous and P. Diaconis, Longest increasing subsequences: From patience sorting to the Baik-Deift-Johansson theorem, preprint.

[4] J. Baik, P. Deift, and K. Johansson, On the distribution of the length of the longest increasing subsequence of random permutations, preprint (math.CO/9810105).

[5] J. Baik, P. Deift, and K. Johansson, On the distribution of the length of the second row of a Young diagram under Plancherel measure, preprint (math.CO/9901118).

[6] J. Baik and E. M. Rains, The asymptotics of monotone subsequences of involutions, preprint (math.CO/9905084).

[7] E. L. Basor, Distribution functions for random variables for ensembles of positive Hermitian matrices, *Commun. Math. Phys.* **188** (1997), 327–350.

[8] E. L. Basor, Connections between random matrices and Szegö limit theorems, in *Spectral Problems in Geometry and Arithmetic*, ed. T. Branson, American Mathematical Society, Providence, 1999, pgs. 1–7.

[9] P. Bleher and A. Its, Semiclassical asymptotics of orthogonal polynomials, Riemann-Hilbert problem, and universality in the matrix model, preprint, Jan. 1997.

[10] O. Bohigas, M. J. Giannoni and C. Schmit, Characterization of chaotic quantum spectra and universality of level fluctuation laws, *Phys. Rev. Lett.* **52** (1984), 1–4.

[11] A. Borodin, Longest increasing subsequences of random colored permutations, *Elect. J. of Combinatorics*, **6**(1) (1999), # R12.

[12] M. J. Bowick and E. Brézin, Universal scaling of the tail of the density of eigenvalues in random matrix models, *Phys. Lett.* **B268** (1991), 21–28.

[13] A. Borodin, A. Okounkov and G. Olshanski, Asymptotics of Plancherel measures for symmetric groups, preprint (math.CO/9905032).

[14] P. A. Deift and X. Zhou, A steepest descent method for oscillatory Riemann-Hilbert problems: Asymptotics for the MKdV equation, *Ann. Math.* **137** (1993), 295–368.

[15] P. A. Deift, A. R. Its and X. Zhou, A Riemann-Hilbert approach to asymtptotic problems arising in the theory of random matrix models, and also in the theory of integrable statistical mechanics, *Ann. of Math.* **146** (1997) 149–235.

[16] P. Deift, T. Kriecherbauer, K. T-R McLaughlin, S. Venakides, and X. Zhou, Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory, preprint.

[17] I. M. Gessel, Symmetric functions and P-recursiveness, *J. Comb. Theory, Ser. A,* **53** (1990), 257–285.

[18] T. Guhr, A. Mueller-Groeling, and H. A. Weidenmueller, Random-matrix theories in quantum physics: common concepts, *Phys. Reports* **299** (1998), 189–425.

[19] J. Harnad, On the bilinear equations for Fredholm determinants appearing in random matrices, preprint (solv-int/9906004).

[20] A. R. Its, A. G. Izergin, V. E. Korepin, and N. A. Slavnov, Differential equations for quantum correlation functions, *Int. J. Mod. Physics* **B4** (1990), 1003–1037.

[21] M. Jimbo, T. Miwa, Y. Môri, and M. Sato, Density matrix of an impenetrable Bose gas and the fifth Painlevé transcendent, *Physica* **1D** (1980), 80–158.

[22] K. Johansson, The longest increasing subsequence in a random permutation and a unitary random matrix model, *Math. Research Lett.* **5** (1998), 63–82.

[23] K. Johansson, Shape fluctuations and random matrices, preprint (math.CO/9903134).

[24] K. Johansson, Discrete orthogonal polynomial ensembles and the Plancherel measure, preprint (math.CO/9906120).

[25] N. M. Katz and P. Sarnak, *Random Matrices, Frobenius Eigenvalues, and Monodromy*, American Mathematical Society, Providence, 1999.

[26] A. B. J. Kuijlaars and K. T-R McLauglin, Generic behavior of the density of states in random matrix theory and equilibrium problems in the presence of real analytic external fields, to appear in *Commun. Pure Appl. Math.*

[27] M. L. Mehta, *Random Matrices*, 2nd ed., Academic Press, San Diego, 1991.

[28] M. L. Mehta, A non-linear differential equation and a Fredholm determinant, *J. de Phys. I France* **2** (1992) 1721–1729.

[29] A. M. Odlyzko, On the distribution of spacings between zeros of the zeta function, *Math. Comp.* **48** (1987), 273–308.

[30] A. M. Odlyzko, B. Poonen, H. Widom and H. S. Wilf, On the distribution of longest increasing subsequences in random permutations, unpublished notes.

[31] A. Okounkov, Random matrices and random permutations, preprint (math.CO/9903176).

[32] J. Palmer, Deformation analysis of matrix models, *Physica D* **78** (1994), 166–185.

[33] E. M. Rains, Increasing subsequences and the classical groups, *Elect. J. of Combinatorics*, **5** (1998), #R12.

[34] A. Regev, Asymptotic values for degrees associated with strips of Young diagrams, *Adv. in Math.* **41** (1981), 115–136.

[35] M. Robnik, Improved energy level statistics for a family of billiards with analytic boundaries, *J. Phys. A* **25** (1992), 3593–3602.

[36] A. Soshnikov, Universality at the edge of the spectrum in Wigner random matrices, preprint (math-ph/9907013).

[37] C. A. Tracy and H. Widom, Correlation functions, cluster functions and spacing distributions for random matrices, *J. Statist. Phys.* **92** (1998), 809–835.

[38] C. A. Tracy and H. Widom, Introduction to random matrices, in *Geometric and Quantum Aspects of Integrable Systems*, G. F. Helminck, ed., Lecture Notes in Physics, Vol. 424, Springer-Verlag (Berlin), 1993, pgs. 103–130.

[39] C. A. Tracy and H. Widom, Level-spacing distributions and the Airy kernel, *Commun. Math. Phys.* **159** (1994), 151–174.

[40] C. A. Tracy and H. Widom, On orthogonal and symplectic ensembles, *Commun. Math. Phys.* **177** (1996), 727–754.

[41] C. A. Tracy and H. Widom, Fredholm determinants, differential equations and matrix models, *Commun. Math. Phys.* **163** (1994), 33–72.

[42] C. A. Tracy and H. Widom, Asymptotics of a class of Fredholm determinants, in *Spectral Problems in Geometry and Arithmetic*, ed. T. Branson, American Mathematical Society, Providence, 1999, pgs. 167–174.

[43] C. A. Tracy and H. Widom, Random unitary matrices, permutations and Painlevé, preprint (math.CO/9811154).

[44] C. A. Tracy and H. Widom, On the distributions of the lengths of the longest monotone subsequences in random words, preprint (math.CO/9904042).

[45] H. Widom, On the relation between orthogonal, symplectic and unitary ensembles, *J. Statist. Phys.* **94** (1999), 347–363.

[46] E. P. Wigner, Distribution of neutron resonance level spacing, Columbia University Report CU–175, 1957. [Reprinted in *The Collected Works of Eugene Paul Wigner*, Part A, Vol. II, Springer-Verlag, Berlin, 1996, pgs. 337–338.]

[47] J. X. Zhong, U. Grimm, R. A. Römer and M. Schreiber, Level-spacing distributions of planar quasiperiodic tight-ginding models, *Phys. Rev. Lett.* **80** (1998), 3996–3999. See also, U. Grimm, R. A. Römer, M. Schreiber, J. X. Zhong, Universal level-spacing statistics in quasiperiodic tight-binding models, preprint (cond-mat/9908063).